

CONFIDENTIAL

SECRET

2

PHASE II FINAL REPORT

PROJECT CHIVE

Volume II

MANAGEMENT SUMMARY

CHIVE/R-3-65

1 March 1965

TOTAL DDC'S HEREIN /

FOLDER NO.

2

BOX NO.

JOB NO.

78-03952A

DIRECTORATE OF SCIENCE AND TECHNOLOGY
OFFICE OF COMPUTER SERVICES

SECRET

GROUP 1
Excluded from automatic
downgrading and
declassification

WARNING

This material contains information affecting the National Defense of the United States within the meaning of the espionage laws, Title 18, USC, Secs. 793 and 794, the transmission or revelation of which in any manner to an unauthorized person is prohibited by law.

~~CONFIDENTIAL~~

U.S. Central Intelligence Agency.
"Office of Computer Services."

Y
151.2
4799
Vol. 2

Project CHIVE:
Phase II Final Report

Volume II

MANAGEMENT SUMMARY

CHIVE/R-3-65

1 March 1965

010365

DQC	1	REV DATE	07/12/81	
ORIG COMP		OPI	63	TYPE
ORIG CLASS	5	PAGES	56	REV CLASS
JUST	22	NEXT REV	2011	AUTH: HR 70-2

25X1

~~CONFIDENTIAL~~

~~SECRET~~ CONFIDENTIAL

TABLE OF CONTENTS

	<u>Page</u>
2.1. System Benefits	1
2.1.1. From the User Point-of-View	1
2.1.2. From the System Operator Point-of-View	3
2.1.3. From the System Manager Point-of-View	5
2.1.4. CHIVE Impact on OCR	6
2.2. Design Implications	9
2.2.1. Long Range Implications	9
2.2.2. Implementation Risks	13
2.3. Design Gaps	17
2.3.1. Security	18
2.3.2. Backup	20
2.3.3. OCR Inherited Files	21
2.3.4. Data Retirement	23
2.3.5. Organization	24

~~SECRET~~ CONFIDENTIAL

~~SECRET~~
CONFIDENTIAL

	<u>Page</u>
2.4. Implementation Alternatives	27
2.4.1. Alternative Roles of the Project Team	27
2.4.2. Operational Alternatives	29
2.4.3. Pre-Implementation Alternatives	32
2.5. Summary Recommendations	37
<hr/>	
2.A. Summary of System Costs	39
2.A.1. Current Costs	39
2.A.2. Implementation Costs	41
2.A.3. Initial Operating Costs	45

~~SECRET~~

CONFIDENTIAL

Approved For Release 2005/06/22 : CIA-RDP78-03952A000100020001-0

TABLES

	<u>Page</u>
2-1 Operational Alternatives	30
2.A-1 Current Costs	40
2.A-2 Implementation Costs	42
2.A-3 Estimated Grade, Rate of Increment and Totals for OCR	43
2.A-4 Initial Operating Costs	46
2.A-5 Cost Estimates for Document Delivery Systems	50, 51

Approved For Release 2005/06/22 : CIA-RDP78-03952A000100020001-0

SECRET

CONFIDENTIAL

~~CONFIDENTIAL~~

Approved For Release 2005/06/21 : CIA-RDP78-03952A000100020001-0

FIGURES

Page

2-1 System Benefits

2

Approved For Release 2005/06/22 : CIA-RDP78-03952A000100020001-0

~~SECRET~~

~~CONFIDENTIAL~~

~~SECRET~~ CONFIDENTIAL

Chapter 2.1.

SYSTEM BENEFITS

This introductory chapter of Volume II forms the bridge between the descriptive material of Volume I and the management implications which follow. Here the anticipated benefits of the system are listed and briefly described. These benefits are summarized in graphic form in Figure 2-1 which shows their relationships and their ultimate focus on a favorable payoff/cost ratio.

2.1.1. FROM THE USER POINT-OF-VIEW

Broad Document Coverage - the system should provide index control over all categories of documents which are of demonstrated value to the system's customers, e.g., open literature (translations), cables, FBIS reports, finished intelligence materials, etc.

High Indexing Specificity - document indexing not by general subject or header control alone, but down to the specific term level

25X1A

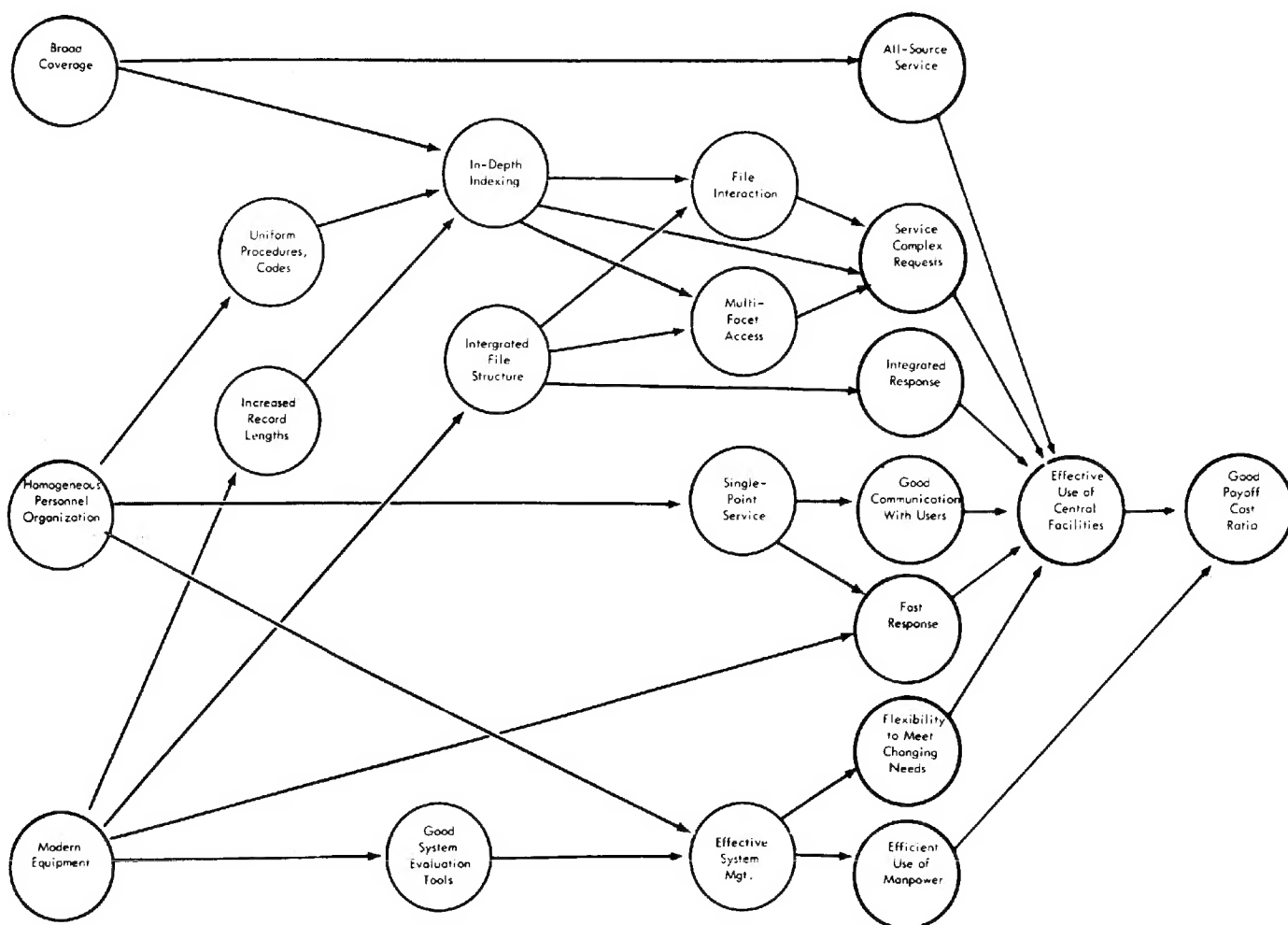
25X1A

and relations between terms.

Exhaustive Indexing - index control of documents by source, area, date, title, general subject content, named object content (persons, installations, commodities)

SYSTEM BENEFITS
User Point-of-View
2.1.1.

~~SECRET~~ CONFIDENTIAL



SECRET

and subjects which are related to or are attributes of named objects.

Capability to Answer Complex Questions - the broad data base and comprehensive indexing, plus the power of the computer to search, collate, compare, relate, rearrange and display data should provide a wide range of query options.

High Retrieval Speeds - modern, high-speed, machine index-search methods combined with rapid document retrieval and reproduction should provide a significant benefit in retrieval speed and selectivity.

Single-Point Service - a customer must contact only one or two system components to obtain a complete answer to his query; a wide range of system files and capabilities would be accessible from any one point of contact.

All-Source Output Capability - the system will combine information from various sources and at various classification levels into an integrated retrieval product.

2.1.2. FROM THE SYSTEM OPERATOR POINT-OF-VIEW

Reduction of Multiple Processing of Identical Documents - one indexer would process a document from the viewpoint of both document and information retrieval, i.e., combined indexing of subjects and named objects. Indexers would work at the direction of information specialists.

SYSTEM BENEFITS

System Operator Point-of-View

SECRET

High Transcription Speeds - reduction of multiple handling reduces multiple coding or transcription to address multiple interests, multiple files. Redundant transcription would be required only for selected summary information files.

Increased Record Length to Reduce File Proliferation - a variable number of data elements can be included in one machine record, obviating the need for many sub-files or "header-trailer" punched card problems.

Limited Manual Labor in Processing System Outputs - the power of the system programs to retrieve, format, and print selected information from index records or from special information files should limit significantly the amount of manual summarization, editing, and typing required in preparation of analyzed system outputs.

Micro-storage Medium - single, compact, high-volume, storage medium would be used for most documents. Easily reproducible. Document never out of file.

Fast Index and Support File Query Response - high system response speeds, together with the ability to produce a wide range of listings of data, should provide the power to respond to short deadlines and increase customer satisfaction.

SYSTEM BENEFITS
System Operator Point-of-View
2.1.2.

SECRET

SECRET

Improved Communications Between Operator and Customer - the all-source environment should give the operator a wider system knowledge to cope with service problems.

Common System Vocabularies - uniformity of system codes and terminology should enable the operator to express his question more precisely to multi-purpose system files and provide an integrated response from all files and sources.

High File Utilization - rapid retrieval and the ability to answer complex queries should produce wide range and large volume of requests. System customers should begin to rely more on the central system.

2.1.3. FROM THE SYSTEM MANAGER POINT-OF-VIEW

Flexibility to Meet Changing Needs - longer record length and specific indexing, plus the selection and formatting power of the computer, should enable the system to meet new requirements without initiation of special programs and without reduction of other services.

Integrated Organizational Structure - integrated system of geographic components enables management to see the system more nearly as a single entity. System uniformity in personnel assignments should allow

SYSTEM BENEFITS
System Manager Point-of-View
2.1.3.

SECRET

~~SECRET~~

flexibility in personnel allocations without the need for continuous retraining.

Good System Evaluation Tools - improved statistical collection mechanisms, together with compatible management reporting formats, should provide better, more easily interpreted information for management evaluation and decision making.

Improved Payoff/Cost Ratio - greater user satisfaction and increased service demands and capabilities should provide a high return for the man/machine investment even if some increase in overall operating cost is required.

2.1.4. CHIVE IMPACT ON OCR

From the foregoing discussion the reader may note an indirect comparison with present OCR procedures or functions. The CHIVE system is designed to encompass most of the functions of the present central reference organization, to reshape to a considerable extent the way these functions are performed, and to upgrade a number of features of the present system which have been recognized as limitations by system operators or customers.

Some of the benefits listed are easily and inexpensively attainable; others are difficult to attain, will be very costly in terms of manpower and funds, and may produce varying degrees of success.

SYSTEM BENEFITS
CHIVE Impact on OCR

~~SECRET~~

SECRET

Their attainment, in the aggregate, could bring about a profound alteration in central reference procedures and organization. Individual functions could be ranked in terms of priority of need, costs, and probably degrees of success (this has been done to a certain extent in Section 2.4.2) and attacked as separate or as related problems.

While benefits are attainable outside the context of CHIVE, the CHIVE design provides a functional integration of this set of benefits and goals. It is in the interrelation of each, in a broad front of implementation, that a long-range gain in terms of effectiveness and payoff vs. cost is most likely to be realized.

SYSTEM BENEFITS
CHIVE Impact on OCR
2.1.4.

SECRET

SECRET

Chapter 2.2.

DESIGN IMPLICATIONS

2.2.1. LONG RANGE IMPLICATIONS

The Agency has held, historically, a position of Community leadership in intelligence information handling dating back to 1947. Recently a responsibility for maintaining this leading position was laid upon the Agency in an action memorandum from the Executive Director-Comptroller who reflected the DCI's directive that the Agency should develop "the best possible system for the handling of information, indexing and data processing."

During the past few years data processing and information handling have become a significant Community preoccupation, and the other intelligence services have spent considerable time and money in development and experimentation. Their activity, to date, has produced many interesting claims and programs, but relatively little in the way of concrete achievement. The Agency has been slower to join in the ADP surge, preferring to engage in extended definition of the problem, rather than yield to pressures to get something "on-the-air" in a hurry. Since its actions have been more deliberate than those of some of the other services, the Agency has made few

DESIGN IMPLICATIONS
Long Range Implications
2.2.1.

SECRET

false starts and has avoided accumulating a record of expensive failures.

While there may be no grounds for the belief that the Agency's leadership in information processing is in imminent jeopardy, there are conversely, no grounds for complacency. State, DIA and NSA are at present vigorously attacking a wide front of information handling problems and (except for State) are heavily supported by experienced manpower and money. In the Agency, volume of incoming documents and technical collection information is steadily increasing, manpower available to analyze, index and store this load has decreased, and the demands on the information services for support are increasing in volume, complexity and required response time. If the Agency is to have the "best possible system for the handling of information," something must be done. This is what CHIVE is all about.

The success of the development of a large-scale, integrated, all-source information handling system in CIA could have a widespread impact on the diverging information handling activities of the Community. The establishment of code and format compatibility standards which could allow for better and more rapid exchange of information between agencies, and the possible

DESIGN IMPLICATIONS
Long Range Implications
2.2.1.

SECRET

elimination of requirements for maintaining files and processing activities which overlap or duplicate those of another agency, have been Community objectives for years, actively and, for years almost uniquely, sought by CIA through the USIB Committee on Documentation (CODIB).

While CHIVE developments would not augur an early impact on standards or compatibility, there is reason to hope that the "service of common concern" coloration which has been with OCR for years could be deepened by the development of new techniques and capabilities which might be emulated by other USIB agencies. In short, the Agency might show by example that a versatile, high-performance system can be built at a reasonable cost. If this is done, external dependence may carry standards and compatibility in its wake.

Within the Agency itself, in addition to the basic objectives of providing more specific and timely responses from the central reference facility, a major system success would open a wide front of implications for more meaningful manipulation of data elements for research, for better support of analyst filing operations, or in some cases, the elimination of the need for extensive analyst files. In long-range terms, it may be possible to foresee some impact on the organization of the Agency production offices

DESIGN IMPLICATIONS
Long Range Implications
2.2.1.

SECRET

~~SECRET~~

through a successful demonstration of the integrated area approach to information processing, storage and retrieval.

As a result of the Phase II work, CHIVE has outlined the design for what it feels to be a realizable, optimum document/information system. This system would be an evolutionary and logical development from an initial system. The implementation stage--the first step from present operations to the initial CHIVE system--involves a radical departure from the mode of operations of the present central reference complex, and requires the interweaving of a multitude of design, training, processing, and analytical functions at a fixed point in the future.

An attempt has been made in Appendix 2.A to estimate the costs of the implementation phase of the project as well as the initial operating costs. While it would be desirable to provide an estimated aggregate cost here for management review, such an estimate would not be meaningful because several variables relating to manning level and equipment alternatives are involved. However, the data available in the tables of Appendix 2.A should be sufficient to estimate aggregate costs when decisions on these variables are made.

The scope of the problem is enormous, and Agency management must judge, as a corollary of its evaluation of the CHIVE Phase II report, what should be the most

DESIGN IMPLICATIONS
Long Range Implications
2.2.1.

~~SECRET~~

SECRET

prudent and effective course to follow. The question that confronts us then appears to be not "Do we proceed?," but "How do we proceed?" and "At what pace do we proceed?". The answers to these questions will probably depend upon management's view of just how critical the Agency's information problem has become--and thus, how urgently we need to get a full-scale operational program underway--and on its assessment of the implementation risks.

2.2.2. IMPLEMENTATION RISKS

There can be no point in minimizing the fact that the initial implementation risks are great. The design gaps and tasks described in this and subsequent volumes of this report will illustrate the magnitude of the problem. Other, unanticipated problems will reveal themselves along the way. The risks are, in a sense, an aggregate of the possible degrees of success listed in the suggested alternative approaches presented in Section 2.4.2. for all functions included in the initial system design.

The danger exists that the initial system could be so beset with problems and handicaps that it will fail from the sheer weight of the technical and human hardships in the implementation environment. But, if the Agency is to retain its Community leadership, the system must not fail. Thus, the basic implementation problem might be

DESIGN IMPLICATIONS
Implementation Risks
2.2.2.

SECRET

SECRET

described as one of creating the most favorable environment in which to begin implementation of an initial CHIVE component.

Management may judge that the information handling problem in the Agency is critical enough to get a live operation under way as rapidly as possible, regardless of the inadequacies or risks. On the other hand, management may judge the climate to be critical, but the implementation risks great enough to argue for adopting as an initial course the automation of a number of related OCR programs or functions, in order to prepare the most favorable climate for the initial system. Some thoughts on possible "pre-implementation" tasks are given in Chapter 2.4.

Suggested below are a few of the questions which management might find useful in making its assessment of the risk factor in the initial implementation plans:

- What aspects of the design appear to offer the highest potential risks?
- Are these aspects critical to success?
- Will it be easy to detect when any basic design hypothesis is in trouble?
- What are the indications of trouble?
- What impact would the failure of any basic design element have on the surrounding design elements?
- How easy will it be to snift direction?

DESIGN IMPLICATIONS
Implementation Risks
2.2.2.

SECRET

SECRET

- Will there be specific, irretrievable losses (beyond development time and money) if the implementation plan must be changed halfway through?
- What are the priorities for attacking major unfilled design gaps?
- What degree of success can be anticipated in filling these gaps?
- Is the scope of implementation too broad?
- What are the possibilities of success in areas where others have failed?

The authors of this report have not attempted to prejudice the case in favor of the proposed system by suggesting answers to these questions, preferring to let the report and the preliminary experimentation speak for themselves. A few observations, however, can be offered for the consideration of those who will attempt to develop answers to these and other questions. First, the design was not laid out with minimum risk as the most important factor; the problem appears to demand success, whatever the risks. Second, it is impossible to critically examine every step to determine whether or not it is safe to take the next one. It is easy to reduce any developmental activity to impotence by continually examining each area of uncertainty; an element of faith--in the design, and in the managers and operators--is required. Finally, the importance of good coordination between designers, users, and operators has been

DESIGN IMPLICATIONS
Implementation Risks
2.2.2.

SECRET

~~SECRET~~

recognized and will be continued to the fullest--even at the expense of what may appear to be a reasonable implementation schedule. No goals will be "declared" until experimentation demonstrates that they have, in fact, been reached.

DESIGN IMPLICATIONS
Implementation Risks
2.2.2.

~~SECRET~~

SECRET

Chapter 2.3.

DESIGN GAPS

The sections below are included to highlight a number of as yet unresolved problems which point to serious gaps remaining in the design.* Several of these are major design problems which were recognized long ago, but which still appear to offer no easy solutions, and some seem not to yield to any solutions which can be derived from combinations of existing and proposed procedures and requirements. It should be noted that most of these would not be problems at all if CHIVE were being developed in an environment where there was no "existing" system. As it is, they arise from the unique design situation of CHIVE--the developing system must be wedded to well-established functions in a well-established agency with all the precedents, procedures, files, regulations and traditions which establishment implies. This is the "bad"--the intellectual and procedural conflict--which must be taken with the "good" of experience that well-established functions can provide.

* Detailed discussion of these problem areas can be found in Volumes V, VI and VII.

DESIGN GAPS
2.3.

SECRET

~~SECRET~~

2.3.1. SECURITY

CHIVE design is especially sensitive to matters of security because of the all-source nature of the data base and because service must be provided to customers cleared for various levels.

CHIVE security at present is a mixture of analyst and computer operator procedures, and of EDP software and hardware processing features. The security level of every document entered in CHIVE is listed in the EDP machine record of the Master Index File, where it is available for use as an automatic search exclusion criterion. The index record provided to the information analyst also contains a printed record of the security level.

The security resolution burden in the all-source files has been placed on the analyst at output time. Alternatively, a diversity of files categorized by security level could be maintained. It has been proposed that the latter course not be followed. However, the trade-off between input and output is an ever-present one, whose validation awaits trial in an initial system.

Analysts or indexers are also required to exercise control over the information data base. The analyst must monitor the security level of records maintained in

DESIGN GAPS
Security
2.3.1.

~~SECRET~~

SECRET

summary information files. He must generate an upgrading transaction before he can enter new data from documents of a more sensitive nature to these files. Denial of information to non-all-source cleared customers because of upgrading of summary files is a problem that must be resolved. When the summary record is upgraded (as mentioned in the above example) the information formerly at a lower level must be denied thereafter to the non-all-source customer who otherwise is entitled to it. The indexer must likewise maintain the security level of entries in the decode file, a file which adds no fundamentally new information content to the data base, but merely transforms codes in index records into clear text which might be at a higher classification level. Protecting the security of this file may cause upgrading of the security levels of many index records.

A large task remains in the security area, since the problem cannot be allowed to develop the operational complexity and potential service limitations which appear to be inherent in the present stage of design. Detailed security procedures will be worked out between OCR and OS.

DESIGN GAPS
Security
2.3.1.

SECRET

~~SECRET~~

2.3.2. BACKUP

CHIVE Backup consists of an integrated set of manual and EDP emergency or loss procedures, together with a plan for duplicating the CHIVE data base. Accidental loss of a file or a system processing element (such as a storage unit) will not result in complete loss of capability in the affected function; only a temporary reduction in operational capability will obtain during the period in which the lost element is being restored. In the case of a damaged EDP file, for example, operations will fall back on a prepared listing; in the case of a storage unit failure, alternate storage units will be utilized.

Unresolved problems exist in the areas of definition of backup requirements and backup scope. No firm performance specifications have been established with respect to mean-time-between-failure, tolerable down time, or failure rate for the various CHIVE hardware components and/or functions. No firm requirements have been established for the need to achieve a crisis backup facility which could sustain operations in an environment in which the on-site installation is lost.

There also exists a requirement for some form of remote backup storage to permit response to specific document requests which cannot be serviced from the

DESIGN GAPS
Backup
2.3.2.

~~SECRET~~

SECRET

headquarters file or to provide a means of reconstruction of the main file in the event of destruction.*

2.3.3. OCR INHERITED FILES

The so-called "inherited" files are the detail index files now maintained by OCR which must be accessed by CHIVE to provide customer service and to produce indexer aids.

2.3.3.1. File Conversion

A basic CHIVE hope has been that the inherited OCR EAM files could be converted to EDP and, where possible, in such a manner as to provide a single-search integrated output. A number of major design decisions have been based on this goal, including the organization configuration developed [] in 1964. At that time, speedy conversion of the SR detail file was recommended. A preliminary examination of the EAM files over the past few months, however, gave rise to doubts about the feasibility of converting and integrating these files. Conversion appeared questionable in terms of cost and obtainable response speeds; integration appeared to be impossible.

25X1A

* For additional comment on these requirements see Volumes V and VI.

DESIGN GAPS
OCR Inherited Files
2.3.3.1.

SECRET

Since implementation decisions hang, in part, on the possibility of arriving at a successful resolution to this problem, a solution is now being vigorously pursued. The picture on the SR conversion appears to be brightening, but other conversion problems remain to be addressed. Unsuccessful resolution may indicate that (a) no integrated system response can be obtained; (b) system response speeds, except in requests for recent information, will be those of the inherited files; (c) analysts must continue to deal with a variety of system codes and outputs; and (d) trained manpower will be required, even after total implementation, to maintain and provide service from inherited files.

2.3.3.2 File Colocation

Relocation of OCR files, or colocation of analyst working files in the proposed geographic divisions, is more a logistical than a system problem. It is important nevertheless, and the course of implementation can be affected by decisions on the efficient disposition of large-volume, inherited document or dossier files.

CHIVE recommends centralization of the major OCR document and EAM files in single locations and the consolidation of each function under single document and single EAM managers. While some files could be colocated with the

DESIGN GAPS
OCR Inherited Files
2.3.3.2.

SECRET

responsible analysts, separation of the large Intellofax or SR document files, by geographic units (if unconverted) seems neither feasible nor desirable. It is understood that space may be a major constraining factor in solutions to the file location problem and that considerable study and flow charting of the potential dislocation factors will be required.

Maintenance of these files at a central point by a trained staff, however, could provide more effective analyst utilization of the developing system as well as the extant central reference system during the entire period of CHIVE's evolution to a total system and after its attainment. Repeated relocation and duplication of files or parts of files would not be required, and movement of documents, files, and analysts could be held to a minimum.

CHIVE does recommend that manual special purpose files, including some small document files of limited interest, and analyst background and snag files be colocated with the CHIVE analysts in the geographic work areas.

2.3.4. DATA RETIREMENT

Various observations and suggestions on the possibilities of retiring files, documents or indexes are

DESIGN GAPS
Data Retirement
2.3.4.

SECRET

~~SECRET~~

scattered throughout the previous CHIVE literature. While this problem has been on the horizon for some time, no direct attack on it has yet been made. In fact, no productive effort can be made on this front until a number of detailed decisions are made on file and program structure and control.

Data retirement requirements have not directly affected the course of implementation up to this point, and are not likely to become a system concern for several years. The ability of the computer system to provide statistical information on file activity and on document use rates will provide a considerable increase in capability over anything now available to OCR to assist in deciding which files or documents should be selected for retirement. File retirement does not imply file destruction. Microimage, tape, or listing storage of inactive files at a remote location should be both economical and easy to maintain. Document media and accessibility problems will be kept in mind as the data retirement problem emerges as well as in planning for the central document delivery system.

2.3.5. ORGANIZATION

The CHIVE design activity is still plagued, to a considerable extent, by doubts about how to organize the

DESIGN GAPS
Organization
2.3.5.

~~SECRET~~

SECRET

input/retrieval activity to best allow for achieving system objectives. As a result of the evaluation of CHIVE's original personnel proposals by the OCR divisions, a modification of the combined indexer/information analyst configuration has been developed for the Phase II report. This latest configuration foresees a processing organization composed of header indexers, area information analysts, and content indexers. This concept has not yet been tested; the results of further analysis of this alternative will have an impact on how to best group the content indexers with their associated information analysts.

To date, CHIVE design has been dominated, perforce, by a preoccupation with input processing. The organizational problem, therefore, must be examined intensely from the standpoint of retrieval requirements; i.e. customer needs and request patterns, and the conclusions derived measured against the apparent efficiency of the man/machine relationship.

It may be sufficient here to list a few additional factors which will bear strongly on the evaluation of the personnel configuration -

- Reduction of duplicative processing/reading (an input requirement) - what are the gains - how do they affect retrieval?

DESIGN GAPS
Organization
2.3.5.

SECRET

- Cost (dollar cost and/or manpower cost) per CHIVE transaction compared with transactions in existing systems.
- Customer satisfaction with products.
- Coordination problems in service and retrieval.
- Input rates balanced against service demands.
- Complexity of input techniques - information analyst/content indexer communications.
- Personnel requirements (clerical vs. analytical--ability of clericals to provide title and expand titles; ability of subprofessionals to perform content indexing).
- Complexity of retrieval techniques - query language, ability of analyst to work effectively with old and new files.
- Retrieval times for document, information requests - gain or loss as compared with present systems.
- Analyst specialization - advantages of various types of specialization.
- Analyst/indexer satisfaction and career development.

The final decisions on the organizational problems that confront the system must be made by OCR. Since OCR is currently re-evaluating its internal organization, it is hoped that CHIVE uncertainties will ultimately be alleviated through the adoption of an organizational configuration which will ease the transition to an EDP-supported system.

DESIGN GAPS
Organization
2.3.5.

~~SECRET~~

Chapter 2.4.

IMPLEMENTATION ALTERNATIVES

Notwithstanding the Agency investment in CHIVE and momentum it has gathered in one direction, the design approach and status are such that adjustments or even major redirections could be made without catastrophic effects. Our approach to system implementation is evolutionary--an "all or nothing" implication should not be drawn from the Phase II recommendations.

2.4.1. ALTERNATIVE ROLES OF THE PROJECT TEAM

Throughout the life of the project there has been continual appraisal by OCS and OCR management of the role which the design team--working in a complex environment of precedent, responsibilities, and operating experience--should assume to insure maximum pay-off to the Agency in its pursuit of sound information processing techniques. Several points-of-view are enumerated here to assist Agency management in providing guidance to the design team through Phase III. Each of these is a significant departure from the plan and mode of operation implied by the remainder of the report.

IMPLEMENTATION ALTERNATIVES
Roles of the Project Team
2.4.1.

~~SECRET~~

~~SECRET~~

As a first alternative, the project could be directed toward the implementation of the basic CHIVE EDP capability which would provide only the mechanical structure for OCR operations. In this mode, the project would capitalize on past CHIVE and OCR experience and on EDP work being done in other intelligence environments, but would no longer be concerned with the data to be provided to and be manipulated by the computer. OCR would be trained in system capabilities and adapt their procedures where they considered the computer to be useful.

Second, emphasis could be placed on optimizing the operations in OCR functional areas (whether or not related to EDP), keeping the long-range CHIVE goals in mind. The document indexing function, for example, might be a candidate for such an approach. Others include document delivery and information file operations. This alternative, which is actually a piece-by-piece implementation, is discussed from a different point of view in the next section.

Third, the punched card operations in OCR could be upgraded to computers. Well organized files could be developed for those portions of its data base which have long-range utility--before the various analytical elements of OCR are reshaped.

IMPLEMENTATION ALTERNATIVES
Roles of the Project Team
2.4.1.

~~SECRET~~

SECRET

Fourth, CHIVE could be a modeling project for existing or proposed OCR systems, perhaps developing and testing a drastically scaled-down version of the initial system proposed in this report, or developing isolated techniques. Such an approach would provide a laboratory environment for the more promising, but more radical techniques developed in Phase II.

Finally, the project team could be recast as an EDP systems staff providing ad hoc services to OCR or other components with heavy data manipulation responsibilities. Techniques developed in Phase II would be inserted in evolving systems at the request of such components as they were proven.

Evaluation of these alternatives is left to the reader, but it should be done in conjunction with a detailed analysis of this report.

2.4.2. OPERATIONAL ALTERNATIVES

A more specific set of alternatives, related directly to segments of the CHIVE system, are shown in Table 2-1, with an estimate given of several evaluative factors--urgency, chance of success, pay-off, and relative cost. The intent here is to break down the proposed system (and some related functions) into

IMPLEMENTATION ALTERNATIVES
Operational Alternatives
2.4.2.

SECRET

Table 2-1

OPERATIONAL ALTERNATIVES

<u>Urgency</u>	<u>Chance of Success</u>	<u>Pay Off If Successful</u>	<u>Relative Cost</u>	<u>Procedural</u>
essential	good	great	medium	integrate and strengthen document indexing
essential	excellent	considerable	medium	consolidate and standardize indexer aids (Vocab. Control Files)
important	good	great	medium	improve analyst work station layout and procedures
desirable	excellent	some	low	improve bibliographic control of documents
desirable	questionable	little	low	improve manual data filing systems
important	good	considerable	high	integrate and upgrade document storage and delivery systems
important	good	considerable	medium	improve coordination between substantive contractor services and OCR reference activities
<u>EDP</u>				
essential	good	great	high	use computer for search of large document index file
desirable	excellent	considerable	medium	generate index aids, dossier index lists, source cards and other reports by computer
desirable	questionable	great	high	convert existing reference files from punched card to computer operations
important	good	considerable	medium	improve data transcription techniques
essential	excellent	considerable	medium	provide EDP data manipulation facility for information files

25X+

SECRET

definable, implementable parts in order to:

--determine which functions might be selected as candidates for implementation

--show the effect of eliminating some functions.

Although the functions can be defined, they are not independent--so they cannot be viewed as a system smorgasbord. The estimates for each function, however, were determined on the basis of independent implementation of each.

Combinatorial possibilities would have an effect on the evaluative factors. For example, the combination of two or more EDP functions would probably lower the aggregate cost, but increase the pay-off and reduce the chance of total success. On the other hand, two or more procedural functions would probably raise the aggregate cost, but increase the pay-off and chance of success. In all the areas listed, the analysis and design is far enough along to permit relatively smooth transition to a sub-set of them.

The discussion of possible design adjustments within each of the functions is provided in detailed discussions scattered throughout the report.

Finally, and most important, the system described in this report is hopefully an optimum, realizable

IMPLEMENTATION ALTERNATIVES
Operational Alternatives
2.4.2.

SECRET

~~SECRET~~

design--it is not an ideal system. Therefore the alternatives suggested in his Section should be viewed as fallback alternatives--from a optimum, realizable system to one of several possible sub-optimum, realizable systems.

2.4.3. PRE-IMPLEMENTATION ALTERNATIVES

What is suggested in the paragraphs below should not be considered an alternative to implementation of the initial CHIVE system; a number of possible alternatives to the initial system are described above (Section 2.4.1.). What is suggested here is a preliminary procedure for removing major obstacles to system success in the event management feels that initial implementation may be premature in the face of the numbers of problems to be addressed, and that maintaining the forward thrust of implementation at its present pace is less critical than the preparation of the most favorable conditions for ultimate system success.

CHIVE has now laid out the basic design and implementation guidelines. With an eye toward these, we could proceed to solve a number of separate but difficult problems related to the overall design. The creation of a number of smaller, related, operational

IMPLEMENTATION ALTERNATIVES
Pre-Implementation Alternatives
2.4.3.

~~SECRET~~

SECRET

programs based on OCR's requirements would serve to inspire confidence in EDP in the potential system users and operators, would delimit the problem areas for the design staff and permit a concentrated attack on each task without dissipation of design skills over too wide a front. It would also provide useful experience in working with EDP input techniques and products, thus providing an initial cadre of experienced and at least semi-trained personnel to work in the initial component. Judicious selection of these pre-implementation tasks could remove major obstacles to system success. There are a number of such tasks which could be attacked immediately.

In the event a decision were made in favor of converting a number of OCR programs to EDP in preparation for eventual implementation of the full initial CHIVE system, OCR, in consultation with OCS, could establish the tasks it wished carried out and set priorities for their completion. Experienced manpower, familiar with the files and programs to be placed under EDP control, should be made available to work directly on these tasks along with members of the design staff. Possible tasks to be accomplished might

IMPLEMENTATION ALTERNATIVES
Pre-Implementation Alternatives
2.4.3.

SECRET

include:

- (a). Test EDP manipulation of large document index files.
- (b). Develop standardized area, organizational, personality, subject/commodity vocabulary control files for document indexing (in conjunction with a. and f.).
- (c). Prepare EDP-based information files on OCR special projects or reports.
- (d). Apply EDP techniques to development of unified management reporting formats and content for OCR current operations and future planning.
- (e). Specify procedures for personality and organization name indexing for document indexers (includes correction and selectivity standards).
- (f). Convert selected OCR EAM files to EDP search (SR and/or Intellofax).
- (g). Prepare EDP-based, formatted summary biographic files and installation summary files on Communist China (or any other area where we could utilize format and capitalize on what was learned during preparation).
- (h). Arrange for receipt of large Library of Congress MIRA file in form which can be machine-searched by author, organization, and subject.
- (i). Develop standardized list of document sources

IMPLEMENTATION ALTERNATIVES
Pre-Implementation Alternatives
2.4.3.

SECRET

dealing with China (or other areas) indicating distribution to OCR components and degrees of processing applied to each source by the various components.

(j). Capture incoming SI teletype information at the source, mechanically reformat and prepare header information for machine search.

(k). Arrange for preparation, indexing and storage in machine-readable form of translated materials on Communist China.

25X1

(l). Convert IPI input, collation and publication activity to EDP.

(m). Convert selected detail manual files to EDP form by stenewriter methods.

It will be seen that any combination of the above could well absorb the greater part of CHIVE's manpower, plus a substantial manpower contribution from OCR, for a year or more.

While work on these EDP or EDP-related tasks was going forward, OCR, again with an eye on the ultimate design goals, could proceed to carry out an internal reorganization of its functions in order to ease the transition from the present system to the CHIVE configuration. OCR is already giving close attention to the possibility of an internal grouping of "like"

IMPLEMENTATION ALTERNATIVES
Pre-Implementation Alternatives
2.4.3.

SECRET

~~SECRET~~

functions, and Volume III of this report discusses some organizational steps which might be taken to assist both OCR and the developing CHIVE activity. In addition to a major overhaul of the document filing and EAM functions, some initial steps might be taken in the Intellofax system to introduce new document categories and to produce the area-oriented topic specialists required by the proposed personnel configuration of the CHIVE system. An all-source document query capability, if not an integrated capability, could be achieved by placing document indexing activities under a single management umbrella.

IMPLEMENTATION ALTERNATIVES
Pre-Implementation Alternatives
2.4.3.

~~SECRET~~

~~SECRET~~

Chapter 2.5.

SUMMARY RECOMMENDATIONS

1. That the DD/I concur in the basic design and implementation plan as outlined in the Phase II report and approve proceeding into Phase III.

2. That OCR establish an initial CHIVE component to test design and implementation concepts.

3. That the initial component be operated in a simulated "live" environment in order to provide a demonstration of system feasibility and capabilities.

4. That the EDP programs required to support the initial CHIVE component and integrate its functions with extant OCR man/machine systems be prepared and coordinated with initial component implementation.

5. That OCS integrate CHIVE requirements for the first two years of Phases III and IV into its overall plans for equipment acquisition and scheduling.

6. That the CHIVE machine processing load in OCS be reviewed after these two years to determine the best ultimate course to follow in the management of the computer element of the system.

7. That the central reference complex prepare for the introduction of the developing CHIVE system through

SUMMARY RECOMMENDATIONS
2.5.

~~SECRET~~

~~SECRET~~

reorganization of functions and selective centralization of files.

8. That OCR study the results of the document delivery analysis and obtain a document image system which will meet CHIVE objectives.

9. That OCR also consider application of EDP techniques, on a special project basis, to current functions which interact with the initial CHIVE component.

10. That the DD/I and DD/S&T allocate manpower and funds required to perform the tasks required under items 2-9 above. (Initial manpower requirements are stated in Vol. III and in the cost summary of this volume).

SUMMARY RECOMMENDATIONS
2.5.

~~SECRET~~

SECRET

Appendix 2.A.

SUMMARY OF SYSTEM COSTS

Cost items discussed in this appendix are classed under current, implementation (anticipated), and initial system costs. In general costs have been broken out as allocated to OCR or OCS. Under each office an additional breakdown by personnel, contract, equipment, and operational costs has been provided. All personnel costs are estimated in man years and a gross dollar value has been assigned each man year figure. Other factors (equipment or operating costs) have been expressed in the form that seemed most reasonable for the estimates which could be obtained, i.e., alternative purchase/rental dollar values for equipment, percentages of CPU time for EDP operations.

2.A.1. CURRENT COSTS

Costs incurred from the beginning of the project to March 1965 are given in Table 2.A-1. Personnel costs, with the exception of contractor costs and Development Divisions's Design Branch costs, have not been budgeted directly for CHIVE but were chargeable to other functions of each office. It should be noted that the considerable OCR manhour investment in repeated designer/contractor briefings and in discussions with task groups, human/

SUMMARY OF SYSTEM COSTS
Current Costs
2.A.1.

SECRET

Table 2.A-1

CURRENT COSTS

OCR 1 February 1964 - 31 March 1965

PERSONNEL:

Man Years

Approximate
Dollar Value

OCR/CHIVE Index Experiment

CHIVE Support Staff

*JOT

Key Punch

OCS April 1962 - 31 March 1965

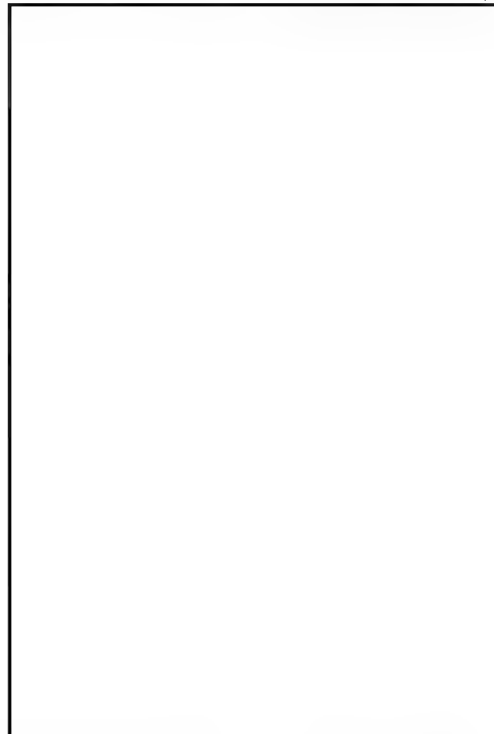
PERSONNEL:

Development Division

Key Punch

Contract Costs:


25X1A



25X1

* One Junior Officer Trainee was assigned to OCR in July 1964 for a Trial Attachment period. Costs were borne by OTR.

25X1A

 dollar costs are rounded off to within 1% of the exact figure; man-year figures are rounded off within 5%.

SECRET

machine time in preparing index experiment tools and in retrieval review, plus OAD document review and meetings are not included in the cost figures for Phase II. DD/I analyst time expended in Phase I and II interviews and fact gathering is also an unaccounted cost item.

2.A.2. IMPLEMENTATION COSTS

Costs given in Table 2.A-2 represent an estimate of those costs which can be anticipated for the 18 month period tentatively projected for completion of the Phase III tasks. Costing, in a number of instances, is somewhat arbitrary, since a number of implementation decisions have not yet been made. Of the OCR costs, the suggested cost of the CHIVE Test Branch is based on a commitment of thirty people. Actual costs will vary with the grades of the persons finally selected for this staff. Estimated grades for these people and the anticipated rates of increment are given in table 2.A-3 below. Since the rate of increment may vary considerably with the rate of progress and availability of skills, this manning level is at best a gross estimate.

Costs given here are based on an estimate of the grades which will probably be transferred with the personnel selected to man the CTB component and partly

SUMMARY OF SYSTEM COSTS
Implementation Costs
2.A.2.

SECRET

IMPLEMENTATION COSTS

PERSONNEL:

CHIVE Support Staff*

EQUIPMENT:

Typewriters (for page reader)

PERSONNEL:

Key Punch



IBM System/360 Model 30 or 40

IBM System/360 Model 60

** Includes 22% beyond base salary for administration and supervision.

**** Because of time sharing on a multi-programmed computer realistic dollar cost estimates are difficult to obtain. Computer useage time is given in terms of estimated total hours running time.

SECRET

on grades held by persons now tentatively designated as a CTB nucleus.

Table 2.A-3

ESTIMATED GRADE, RATE OF INCREMENT AND TOTALS FOR OCR

<u>GS grade & number of personnel</u>		<u>Duration of participation</u>
CSS:	*	
CTB:	<div></div>	18 months
		15 months
		9 months

*Note: No figures are included here for the CHIVE Support Staff and no costs are charged to CHIVE in the charts which follow. Subsequent to the drafting of this chapter the OCR System Analysis Staff was formed and will carry out some of the functions of the former CSS. It has not been determined what percentage of SAS's time will be available to CHIVE. No SAS costs, therefore, have been included for Phase III.

Manpower for the Document Delivery System is based on an average salary for persons performing similar tasks; it averages out to equal GS-4 or 5 ratings.

SUMMARY OF SYSTEM COSTS
Implementation Costs
2.A.2.

SECRET

~~SECRET~~

Space costs have not been included as it is assumed that extant space in the Headquarters building will be allotted to CHIVE. It is also assumed that the Office of Logistics will bear reasonable costs to modify space for the Document Delivery System or to secure work areas not now behind the SI barrier. Some planning and reconstruction toward the latter of these tasks is already under way in OCR.

No estimates for equipment and supplies, courier service, telephones and other communication devices (tube modifications) have been included as the requirement for these is not yet firm enough to permit a reasonable estimate.

The basic assumptions for the costs of the Document Delivery System are given below. These assumptions also apply to the initial operating costs.

- Equipment is to be purchased, wherever possible, provided it is economically advantageous to do so.
- The capital expenditure for the initial system equipment will be made during the Implementation Period (Phase III).
- The operating costs (i.e., personnel, materials, equipment rental and maintenance) are projected for a ten-month period in Phase III. Installation of all equipment is projected by the beginning of Implementation Month 9.

SUMMARY OF SYSTEM COSTS
Implementation Costs
2.A.2.

~~SECRET~~

~~SECRET~~
CONFIDENTIAL

- Materials costs are based on calculated unit costs extended over a ten-month period with 3% added for spoilage.
- Personnel costs are based on observed average annual salary scaled for specific job categories. The annual rates are scaled down to a ten-month period and a 22% additional cost is added to approximate the cost of administration, supervision, leave and training.

EDP equipment usage costs (operating costs in hours) in Table 2.A-2 are based on the assumption that (a) 200 man months of programming effort will be expended over a 15-month period (b) the equipment configuration available from Month 3 through Month 11 will include disk storage and (c) during the latter portion of the 18-month period (after Month 12) testing will be done in a multi-programmed system.

2.A.3. INITIAL OPERATING COSTS

Table 2.A-4 shows the cost estimates for the initial operating period. The cost assumptions used for the initial system are the same as those applied to the anticipated costs except that the projection is given for all functions and personnel for a full 12-month period.

The estimated usage of the 360/60 Central Processor (the most critical cost element of the EDP System) was derived by dividing the processing into five major functional areas: searching, file maintenance, hit

SUMMARY OF SYSTEM COSTS
Initial Operating Costs
2.A.3.

~~SECRET~~

CONFIDENTIAL

~~SECRET~~
CONFIDENTIAL

Table 2.A-4

INITIAL OPERATING COSTS

OCR (one year)

PERSONNEL:

Man Years

Approximate
Dollar Value

China Test Branch

CHIVE Support Staff

Document Delivery System

Ap. Card

Microfiche

EQUIPMENT:

Document Delivery System

Ap. Card - purchased

Ap. Card - rental

Microfiche - purchased

Microfiche - rental

OCS (one year)

PERSONNEL:

Development Division

Contractor

EQUIPMENT:

2060 Central Processor with bulk core

1403 Printer

1402 Reader-Punch

2400 Magnetic Tapes (8)

1302 Disc Storage

45 million characters - permanent

10 million characters - temporary

Console

Page Reader

15% *

13%

1%

5%

13%

100%

**

* Because of time sharing on a multi-programmed computer, realistic cost estimates are difficult to obtain. Percentages are based on 176 hour (one shift) operation and are estimates of use rate at the end of one year.

** Depends on make and model chosen. This data is not yet available.

25X1

~~SECRET~~
CONFIDENTIAL

~~SECRET~~

~~CONFIDENTIAL~~

processing, sorting, and printing. The basic program loops for each function were timed and then multiplied by the number of executions for each loop. It is important to note that the number of executions have been projected on several assumptions. These assumptions that have the greatest impact on the estimate are the number of documents indexed, the number of search requests levied on the machine files and the number of individual searches triggered by each request.

Searching - The search time was estimated as follows:

100 requests per day
10 terms/request
15 compares/record
60,000 records
35 microsecond/compare

The "compare" time estimate is based on an average length per search argument of 10 characters. The class of search operands is greater than, less than, equal, or not equal. Therefore from the above:

(35 microseconds/compare) (15 compares/record)
(100 requests/day) ($6 \cdot 10^4$ records) or, approximately 1 hour/day

File Maintenance - Maintenance processing time was estimated as follows:

25 terms/record
680 characters/record
300 records/day
5,000 microseconds/record

SUMMARY OF SYSTEM COSTS
Initial Operating Costs
2 A.3.

~~CONFIDENTIAL~~

~~SECRET~~ ~~CONFIDENTIAL~~

From the above:

(300 records/day) (500 microseconds/record) or,
approximately 1 minute/day

Hit Processing - Hit processing is the manipulation performed on the records that have been retrieved by the search mechanism. The functions considered are similar to those performed in file maintenance:

100 requests/day
40 drops/request

From the above:

(4000 records/day) (5000 microseconds/record) or,
approximately 1 minute/day

Sorting - Sort requirements include the daily search drops and the maintenance input. The times were supplied by the equipment manufacturer. Estimate:

4000 drop records	4 minutes
300 header records	
300 content records	1 minute

Printing - The basic print load was considered to be supplied by the retrieval process. The processing was estimated as follows:

15 lines to display a record
Decoding
Automatic top/bottom captions
4000 records/day
6500 microseconds/record

From the above:

(4000 records/day) (6500 microseconds/records) or,
approximately 1 minute/day

SUMMARY OF SYSTEM COSTS
Initial Operating Costs

~~CONFIDENTIAL~~

~~SECRET~~

CONFIDENTIAL

Summary - The total use rate of the central processor was derived as follows:

Searching	60 min./day
File Maintenance	1 "
Hit Processing	1 "
Sorting	5 "
Printing	<u>1</u> "
	68 "

$\frac{68}{480}$ or, approximately 15% of an 8 hour shift.

Back-up data on the cost figures given earlier for the alternate document delivery systems are given in Table 2.A-5. These data are based on the assumptions stated in Section 2.A.2.

SUMMARY OF SYSTEM COSTS
Initial Operating Costs
2.A.3.

CONFIDENTIAL

CONFIDENTIAL

Approved For Release 2005/06/22 : CIA-RDP78-03952A000100020001-0

Table 2.A-5

COST ESTIMATES FOR DOCUMENT DELIVERY SYSTEMS

Implementation Period (Delivery and Checkout -- 10 Month Period)

PACKED MICROFICHE SYSTEM

<u>Equipment Cost</u>	<u>Purchase</u>	<u>10 Mos. Rental</u>	<u>10 Mos. Maintenance</u>
1 - Camera		\$ 7,500	
1 - Typewriter *	440		
1 - Film Processor **	12,000		
1 - Ozalid Printer	7,500		250
1 - Film Backer	5,000		
1 - Card File	4,680		85
1 - Film Cutter	2,024		15
1 - Microfiche Printer	48,000		6,650
	<u>\$ 79,644</u>	<u>\$ 7,500</u>	<u>\$ 7,900</u>
<u>Materials</u> (for 10 months including 3% spoilage)			\$ 10,085

Personnel

1 - Camera Operator	@ 4580 per yr.
1 - File Clerk	@ 4690 per yr.
1 - Photo Technician	@ 4690 per yr.
	<u>13,960 per yr.</u>

Total personnel cost plus 22% for administration and supervision: \$14,135

FILMSORT 2000d X APERTURE CARD SYSTEM

<u>Equipment Cost</u>	<u>Purchase</u>	<u>10 Mos. Rental</u>	<u>10 Mos. Maintenance</u>
1 - Camera	\$ 5,090		\$ 280
1 - Key punch (026) ***		\$ 600	

* There are a number of these units on board which could probably be used.

** There is a 105mm developing capability Also there is a basic Pako 17-2 Processing Unit currently being installed which would require about a \$500 change for 105mm.

*** There are a number of these units currently installed which could probably be used during the early implementation phases.

SECRET

CONFIDENTIAL

25X1A

~~SECRET~~ CONFIDENTIAL

Table 2.A-5 (cont.)

<u>Equipment Cost</u>	<u>Purchase</u>	<u>10 Mos. Rental</u>	<u>10 Mos. Maintenance</u>
1 - Verifier (056) *		\$ 500	
1 - Interpreter (557) *		1,650	
2 - Card Files (motorized)	\$ 9,360		\$ 160
1 - Quadrant Printer	1,795		160
1 - Copy Reproducer	<u>65,000</u> \$81,245	<u>\$2,750</u>	<u>1,410</u> \$2,010
<u>Materials</u>	(for 10 months including 3% spoilage)		\$20,520
<u>Personnel</u>			
1 - Camera Operator (1st shift)		@ 4580 per yr.	
1 - Camera Operator (2nd shift)		@ 5267 per yr.	
1 - File Clerk		@ 4690 per yr.	
1 - Photo Technician		@ 5010 per yr. 19,547 per yr. (16,225 per 10 months)	

Total personnel cost plus 22% for administration and supervision: \$19,800

<u>First Year's Operating Cost</u>	<u>Microfiche</u>	<u>Filmsort - Aperture Card</u>
Personnel Cost (inc. 22%)	\$ 15,030	\$ 23,850
Equipment Rental	9,000	3,300 **
Equipment Maintenance	8,420	3,340
Material Cost (inc. 3% waste)	<u>12,150</u> \$ 44,600	<u>24,725</u> \$ 55,215
Space Required (sq. ft.) (inc. 30% misc. space)	675	510

* There are a number of these units currently installed which could probably be used during the early implementation phases.

** This figure (\$3,300) represents EAM equipment equivalent to units currently installed.

~~SECRET~~ CONFIDENTIAL

SECRET

SECRET